

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-032100

(43)Date of publication of application : 31.01.2002

(51)Int.Cl. G10L 19/00
G10L 19/02
H03M 7/36

(21)Application number : 2001-150411 (71)Applicant : LUCENT TECHNOLOG INC

(22)Date of filing : 21.05.2001 (72)Inventor : GHITZA ODED

(30)Priority

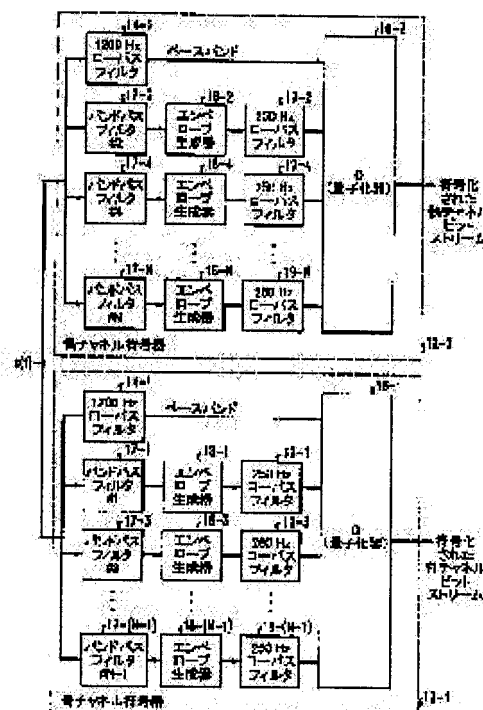
Priority number : 2000 580215 Priority date : 26.05.2000 Priority country : US

(54) METHOD FOR ENCODING AUDIO SIGNAL

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a single non-hybrid type audio encoding system which operates excellently to both of a speech signal and music signal, using characteristics of human aural system.

SOLUTION: A source signal to be encoded is divided into plural frequency bands, and concerning the frequencies lower than a given threshold frequency, the waveforms are encoded so as to store the phase information, however, concerning the frequencies higher than the threshold value, the waveform of each frequency band of these (high frequency side) is decomposed into a carrier signal and an envelope signal, and (not phase information of the carrier signal, but) only the envelope signal is encoded. Thereafter, a decoder decodes these encoded envelope signals and the decoded signals are use for modulating the corresponding carrier waveforms. The carrier waveform is, for example, a constant cosine carrier having a frequency corresponding to a center frequency of a frequency band.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2002-32100

(P2002-32100A)

(43) 公開日 平成14年1月31日 (2002.1.31)

(51) Int.Cl. ⁷	識別記号	F I	テーマコード* (参考)
G 1 0 L 19/00		H 0 3 M 7/36	5 D 0 4 5
	19/02	G 1 0 L 9/18	M 5 J 0 6 4
H 0 3 M 7/36		7/04	G

審査請求 未請求 請求項の数60 O L (全 13 頁)

(21) 出願番号 特願2001-150411(P2001-150411)
(22) 出願日 平成13年5月21日 (2001.5.21)
(31) 優先権主張番号 09/580215
(32) 優先日 平成12年5月26日 (2000.5.26)
(33) 優先権主張国 米国 (U S)

(71) 出願人 596077259
ルーセント テクノロジーズ インコーポ
レイテッド
Lucent Technologies
Inc.
アメリカ合衆国 07974 ニュージャージ
ー、マレーヒル、マウンテン アベニュー
600-700
(74) 代理人 100081053
弁理士 三俣 弘文

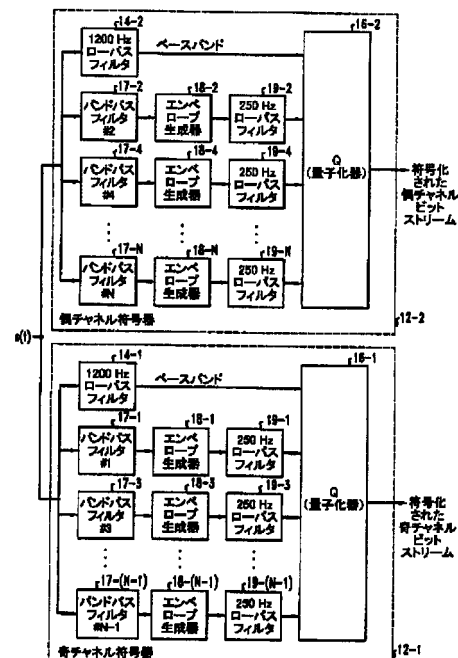
最終頁に続く

(54) 【発明の名称】 オーディオ信号を符号化する方法

(57) 【要約】

【課題】 人間聴覚系の性質を利用して、音声および音楽の両方の信号に対して良好に動作する単一の非ハイブリッド型オーディオ符号化方式を実現する。

【解決手段】 符号化されるべきソース信号は複数の周波数帯域に分割され、与えられたしきい値周波数より低い周波数については、波形は、その位相情報を保存するように符号化されるが、そのしきい値より高い周波数については、これらの（高周波側の）それぞれの周波数帯域の波形は、キャリア信号とエンベロープ信号に分解され、（キャリア信号の位相情報ではなく）エンベロープ信号のみが符号化される。その後、復号器では、これらの符号化されたエンベロープ信号が復号され、対応するキャリア波形を変調するために用いられる。キャリア波形は、例えば、周波数帯域の中心周波数に対応する周波数を有する一定の余弦キャリアである。



【特許請求の範囲】

【請求項1】 オーディオ信号を符号化する方法において、

(a) オーディオ信号を、与えられたしきい値周波数より低い周波数成分を含む低周波数帯域信号と、該与えられたしきい値周波数より高い周波数成分を含む高周波数帯域信号とを含む複数の周波数帯域信号とに分割するステップと、

(b) 前記低周波数帯域信号を表す波形に含まれる位相情報の少なくとも一部を保存するように前記低周波数帯域信号の少なくとも1つを符号化するステップと、

(c) 前記高周波数帯域信号の少なくとも1つに対して、該高周波数帯域信号を表す波形のエンベロープの少なくとも一部を表すが該高周波数帯域信号を表す波形に関連する位相情報を実質的に含まない、対応する臨界帯域エンベロープ信号を生成するステップと、

(d) 対応する臨界帯域エンベロープ信号を符号化することによって前記高周波数帯域信号の少なくとも一部を符号化するステップとを有することを特徴とする、オーディオ信号を符号化する方法。

【請求項2】 前記オーディオ信号は音声信号であることを特徴とする請求項1記載の方法。

【請求項3】 前記オーディオ信号は音楽信号であることを特徴とする請求項1記載の方法。

【請求項4】 前記周波数帯域は、バークスケールに沿ってほぼ一様に分布することを特徴とする請求項1記載の方法。

【請求項5】 前記分割するステップは、ガンマトーンフィルタを用いて実行されることを特徴とする請求項1記載の方法。

【請求項6】 前記与えられたしきい値周波数は約1200Hzであることを特徴とする請求項1記載の方法。

【請求項7】 それぞれの臨界帯域エンベロープ信号をローパスフィルタリングするステップをさらに有し、前記高周波数帯域信号の少なくとも一部を符号化するステップは、対応するフィルタリングされた臨界帯域エンベロープ信号を符号化することを特徴とする請求項1記載の方法。

【請求項8】 前記臨界帯域エンベロープ信号をローパスフィルタリングするステップは、約250Hzのローパスフィルタを用いて実行されることを特徴とする請求項7記載の方法。

【請求項9】 符号化された高周波数帯域信号の第1の真部分集合からなる第1符号化ビットストリームと、符号化された高周波数帯域信号の第2の真部分集合からなる第2符号化ビットストリームとを生成するステップをさらに有し、

前記第1と第2の真部分集合は、互いに素であり、合わせて、前記高周波数帯域信号のすべてを構成することを特徴とする請求項1記載の方法。

【請求項10】 前記高周波数帯域信号を周波数増大順に並べた場合、前記第1および第2の真部分集合は、1つおきの高周波数帯域信号を含むことを特徴とする請求項9記載の方法。

【請求項11】 (a) オーディオ信号を、与えられたしきい値周波数より低い周波数成分を含む低周波数帯域信号と、該与えられたしきい値周波数より高い周波数成分を含む高周波数帯域信号とを含む複数の周波数帯域信号とに分割するステップと、(b) 前記低周波数帯域信号を表す波形に含まれる位相情報の少なくとも一部を保存するように前記低周波数帯域信号の少なくとも1つを符号化するステップと、(c) 前記高周波数帯域信号の少なくとも1つに対して、該高周波数帯域信号を表す波形のエンベロープの少なくとも一部を表すが該高周波数帯域信号を表す波形に関連する位相情報を実質的に含まない、対応する臨界帯域エンベロープ信号を生成するステップと、(d) 対応する臨界帯域エンベロープ信号を符号化することによって前記高周波数帯域信号の少なくとも一部を符号化するステップと、を実行することによって符号化されたオーディオ信号を復号する方法において、

(i) それぞれの符号化された低周波数帯域信号を復号して、対応する低周波数帯域信号を表す再構成波形を生成するステップと、

(ii) それぞれの符号化された高周波数帯域信号を復号して、該高周波数帯域信号に対応する再構成臨界帯域エンベロープ信号を生成するステップと、

(iii) 前記再構成臨界帯域エンベロープ信号のそれぞれを、対応するキャリア波形と結合して、対応する高周波数帯域信号を表す再構成波形を生成するステップと、

(iv) 前記対応する低周波数帯域信号を表す再構成波形と前記対応する高周波数帯域信号を表す再構成波形とをそれぞれ結合して、符号化されたオーディオ信号を表す再構成波形を生成するステップとを有することを特徴とする、オーディオ信号を復号する方法。

【請求項12】 前記オーディオ信号は音声信号であることを特徴とする請求項11記載の方法。

【請求項13】 前記オーディオ信号は音楽信号であることを特徴とする請求項11記載の方法。

【請求項14】 前記周波数帯域は、バークスケールに沿ってほぼ一様に分布することを特徴とする請求項11記載の方法。

【請求項15】 前記与えられたしきい値周波数は約1200Hzであることを特徴とする請求項11記載の方法。

【請求項16】 前記オーディオ信号の符号化は、符号化された高周波数帯域信号の第1の真部分集合からなる第1符号化ビットストリームと、符号化された高周波数帯域信号の第2の真部分集合からなる第2符号化ビット

ストリームとを生成するステップを実行することをさらに含み、

前記第1と第2の真部分集合は、互いに素であり、合わせて、前記高周波数帯域信号のすべてを構成し、

前記復号する方法は、前記第1符号化ビットストリームと前記第2符号化ビットストリームのそれぞれに対して独立に実行されることにより、第1再構成波形および第2再構成波形を生成し、それぞれの再構成波形は、符号化されたオーディオ信号を表すことを特徴とする請求項11記載の方法。

【請求項17】 前記高周波数帯域信号を表す再構成波形を生成するステップは、それぞれのキャリア波形を、対応する再構成臨界帯域エンベロープ信号で変調することを含むことを特徴とする請求項11記載の方法。

【請求項18】 それぞれのキャリア波形は、対応する固定周波数の余弦波系を有することを特徴とする請求項17記載の方法。

【請求項19】 前記キャリア波形の固定周波数は、対応する高周波数帯域信号の中心周波数にほぼ等しいことを特徴とする請求項18記載の方法。

【請求項20】 高周波数帯域信号を表すそれぞれの再構成波形を、対応するノイズ信号と結合するステップをさらに有することを特徴とする請求項11記載の方法。

【請求項21】 オーディオ信号を符号化する装置において、

(a) オーディオ信号を、与えられたしきい値周波数より低い周波数成分を含む低周波数帯域信号と、該与えられたしきい値周波数より高い周波数成分を含む高周波数帯域信号とを含む複数の周波数帯域信号とに分割する手段と、

(b) 前記低周波数帯域信号を表す波形に含まれる位相情報の少なくとも一部を保存するように前記低周波数帯域信号の少なくとも1つを符号化する手段と、

(c) 前記高周波数帯域信号の少なくとも1つに対して、該高周波数帯域信号を表す波形のエンベロープの少なくとも一部を表すが該高周波数帯域信号を表す波形に関連する位相情報を実質的に含まない、対応する臨界帯域エンベロープ信号を生成する手段と、

(d) 対応する臨界帯域エンベロープ信号を符号化することによって前記高周波数帯域信号の少なくとも一部を符号化する手段とを有することを特徴とする、オーディオ信号を符号化する装置。

【請求項22】 前記オーディオ信号は音声信号であることを特徴とする請求項21記載の装置。

【請求項23】 前記オーディオ信号は音楽信号であることを特徴とする請求項21記載の装置。

【請求項24】 前記周波数帯域は、バークスケールに沿ってほぼ一様に分布することを特徴とする請求項21記載の装置。

【請求項25】 前記分割する手段は、ガンマトーンフ

ィルタを有することを特徴とする請求項21記載の装置。

【請求項26】 前記与えられたしきい値周波数は約1200Hzであることを特徴とする請求項21記載の装置。

【請求項27】 それぞれの臨界帯域エンベロープ信号をローパスフィルタリングする手段をさらに有し、前記高周波数帯域信号の少なくとも一部を符号化する手段は、対応するフィルタリングされた臨界帯域エンベロープ信号を符号化する手段を含むことを特徴とする請求項21記載の装置。

【請求項28】 前記臨界帯域エンベロープ信号をローパスフィルタリングする手段は、約250Hzのローパスフィルタを有することを特徴とする請求項27記載の装置。

【請求項29】 符号化された高周波数帯域信号の第1の真部分集合からなる第1符号化ビットストリームと、符号化された高周波数帯域信号の第2の真部分集合からなる第2符号化ビットストリームとを生成する手段をさらに有し、

前記第1と第2の真部分集合は、互いに素であり、合わせて、前記高周波数帯域信号のすべてを構成することを特徴とする請求項21記載の装置。

【請求項30】 前記高周波数帯域信号を周波数増大順に並べた場合、前記第1および第2の真部分集合は、1つおきの高周波数帯域信号を含むことを特徴とする請求項29記載の装置。

【請求項31】 (a) オーディオ信号を、与えられたしきい値周波数より低い周波数成分を含む低周波数帯域信号と、該与えられたしきい値周波数より高い周波数成分を含む高周波数帯域信号とを含む複数の周波数帯域信号とに分割するステップと、(b) 前記低周波数帯域信号を表す波形に含まれる位相情報の少なくとも一部を保存するように前記低周波数帯域信号の少なくとも1つを符号化するステップと、(c) 前記高周波数帯域信号の少なくとも1つに対して、該高周波数帯域信号を表す波形のエンベロープの少なくとも一部を表すが該高周波数帯域信号を表す波形に関連する位相情報を実質的に含まない、対応する臨界帯域エンベロープ信号を生成するステップと、(d) 対応する臨界帯域エンベロープ信号を符号化することによって前記高周波数帯域信号の少なくとも一部を符号化するステップと、を実行することによって符号化されたオーディオ信号を復号する装置において、

(i) それぞれの符号化された低周波数帯域信号を復号して、対応する低周波数帯域信号を表す再構成波形を生成する手段と、

(ii) それぞれの符号化された高周波数帯域信号を復号して、該高周波数帯域信号に対応する再構成臨界帯域エンベロープ信号を生成する手段と、

(i i i) 前記再構成臨界帯域エンベロープ信号のそれぞれを、対応するキャリア波形と結合して、対応する高周波数帯域信号を表す再構成波形を生成する手段と、

(i v) 前記対応する低周波数帯域信号を表す再構成波形と前記対応する高周波数帯域信号を表す再構成波形とをそれぞれ結合して、符号化されたオーディオ信号を表す再構成波形を生成する手段とを有することを特徴とする、オーディオ信号を復号する装置。

【請求項32】 前記オーディオ信号は音声信号であることを特徴とする請求項31記載の装置。

【請求項33】 前記オーディオ信号は音楽信号であることを特徴とする請求項31記載の装置。

【請求項34】 前記周波数帯域は、バークスケールに沿ってほぼ一様に分布することを特徴とする請求項31記載の装置。

【請求項35】 前記与えられたしきい値周波数は約1200Hzであることを特徴とする請求項31記載の装置。

【請求項36】 前記オーディオ信号の符号化は、符号化された高周波数帯域信号の第1の真部分集合からなる第1符号化ビットストリームと、符号化された高周波数帯域信号の第2の真部分集合からなる第2符号化ビットストリームとを生成するステップを実行することをさらに含み、

前記第1と第2の真部分集合は、互いに素であり、合わせて、前記高周波数帯域信号のすべてを構成し、

前記復号する装置は、前記第1符号化ビットストリームと前記第2符号化ビットストリームのそれぞれを独立に復号することにより第1再構成波形および第2再構成波形を生成する手段を有し、それぞれの再構成波形は、符号化されたオーディオ信号を表すことを特徴とする請求項31記載の装置。

【請求項37】 前記高周波数帯域信号を表す再構成波形を生成する手段は、それぞれのキャリア波形を、対応する再構成臨界帯域エンベロープ信号で変調する手段を含むことを特徴とする請求項36記載の装置。

【請求項38】 それぞれのキャリア波形は、対応する固定周波数の余弦波系を有することを特徴とする請求項37記載の装置。

【請求項39】 前記キャリア波形の固定周波数は、対応する高周波数帯域信号の中心周波数にほぼ等しいことを特徴とする請求項38記載の装置。

【請求項40】 高周波数帯域信号を表すそれぞれの再構成波形を、対応するノイズ信号と結合する手段をさらに有することを特徴とする請求項31記載の装置。

【請求項41】 オーディオ信号を符号化する装置において、

(a) オーディオ信号を、与えられたしきい値周波数より低い周波数成分を含む低周波数帯域信号と、該与えられたしきい値周波数より高い周波数成分を含む高周波数

帯域信号とを含む複数の周波数帯域信号とに分割するフィルタバンクと、

(b) 前記低周波数帯域信号を表す波形に含まれる位相情報の少なくとも一部を保存するように前記低周波数帯域信号の少なくとも1つを符号化する符号器と、

(c) 高周波数帯域信号の1つに対して、該高周波数帯域信号を表す波形のエンベロープの少なくとも一部を表すが該高周波数帯域信号を表す波形に関連する位相情報を実質的に含まない、対応する臨界帯域エンベロープ信号を生成するエンベロープ生成器と、

(d) 対応する臨界帯域エンベロープ信号を符号化することによって前記高周波数帯域信号の少なくとも一部を符号化する符号器とを有することを特徴とする、オーディオ信号を符号化する装置。

【請求項42】 前記オーディオ信号は音声信号であることを特徴とする請求項41記載の装置。

【請求項43】 前記オーディオ信号は音楽信号であることを特徴とする請求項41記載の装置。

【請求項44】 前記周波数帯域は、バークスケールに沿ってほぼ一様に分布することを特徴とする請求項41記載の装置。

【請求項45】 前記フィルタバンクは、ガンマトーンフィルタを有することを特徴とする請求項41記載の装置。

【請求項46】 前記与えられたしきい値周波数は約1200Hzであることを特徴とする請求項41記載の装置。

【請求項47】 対応する臨界帯域エンベロープ信号に適用されるローパスフィルタをさらに有し、

前記高周波数帯域信号の少なくとも一部を符号化する符号器は、対応するフィルタリングされた臨界帯域エンベロープ信号を符号化する符号器を含むことを特徴とする請求項41記載の装置。

【請求項48】 前記ローパスフィルタは、約250Hzのローパスフィルタを有することを特徴とする請求項47記載の装置。

【請求項49】 符号化された高周波数帯域信号の第1の真部分集合からなる第1符号化ビットストリームと、符号化された高周波数帯域信号の第2の真部分集合からなる第2符号化ビットストリームとが生成され、

前記第1と第2の真部分集合は、互いに素であり、合わせて、前記高周波数帯域信号のすべてを構成することを特徴とする請求項41記載の装置。

【請求項50】 前記高周波数帯域信号を周波数増大順に並べた場合、前記第1および第2の真部分集合は、1つおきの高周波数帯域信号を含むことを特徴とする請求項49記載の装置。

【請求項51】 (a) オーディオ信号を、与えられたしきい値周波数より低い周波数成分を含む低周波数帯域信号と、該与えられたしきい値周波数より高い周波数成分

分を含む高周波数帯域信号とを含む複数の周波数帯域信号とに分割するステップと、(b)前記低周波数帯域信号を表す波形に含まれる位相情報の少なくとも一部を保存するように前記低周波数帯域信号の少なくとも1つを符号化するステップと、(c)前記高周波数帯域信号の少なくとも1つに対して、該高周波数帯域信号を表す波形のエンベロープの少なくとも一部を表すが該高周波数帯域信号を表す波形に関連する位相情報を実質的に含まない、対応する臨界帯域エンベロープ信号を生成するステップと、(d)対応する臨界帯域エンベロープ信号を符号化することによって前記高周波数帯域信号の少なくとも一部を符号化するステップと、を実行することによって符号化されたオーディオ信号を復号する装置において、

(i)それぞれの符号化された低周波数帯域信号を復号して、対応する低周波数帯域信号を表す再構成波形を生成する復号器と、

(i i)それぞれの符号化された高周波数帯域信号を復号して、該高周波数帯域信号に対応する再構成臨界帯域エンベロープ信号を生成する復号器と、

(i i i)前記再構成臨界帯域エンベロープ信号のそれぞれを、対応するキャリア波形と結合して、対応する高周波数帯域信号を表す再構成波形を生成する乗算器と、

(i v)前記対応する低周波数帯域信号を表す再構成波形と前記対応する高周波数帯域信号を表す再構成波形とをそれぞれ結合して、符号化されたオーディオ信号を表す再構成波形を生成する加算器とを有することを特徴とする、オーディオ信号を復号する装置。

【請求項52】 前記オーディオ信号は音声信号であることを特徴とする請求項51記載の装置。

【請求項53】 前記オーディオ信号は音楽信号であることを特徴とする請求項51記載の装置。

【請求項54】 前記周波数帯域は、バークスケールに沿ってほぼ一様に分布することを特徴とする請求項51記載の装置。

【請求項55】 前記与えられたしきい値周波数は約1200Hzであることを特徴とする請求項51記載の装置。

【請求項56】 前記オーディオ信号の符号化は、符号化された高周波数帯域信号の第1の真部分集合からなる第1符号化ビットストリームと、符号化された高周波数帯域信号の第2の真部分集合からなる第2符号化ビットストリームとを生成するステップを実行することをさらに含み、

前記第1と第2の真部分集合は、互いに素であり、合わせて、前記高周波数帯域信号のすべてを構成し、

前記復号する装置は、前記第1符号化ビットストリームと前記第2符号化ビットストリームのそれぞれを独立に復号することにより第1再構成波形および第2再構成波形を生成し、それぞれの再構成波形は、符号化されたオ

ーディオ信号を表すことを特徴とする請求項51記載の装置。

【請求項57】 前記乗算器は、対応するキャリア波形を、対応する再構成臨界帯域エンベロープ信号で変調することを特徴とする請求項56記載の装置。

【請求項58】 それぞれのキャリア波形は、対応する固定周波数の余弦波系を有することを特徴とする請求項57記載の装置。

【請求項59】 前記キャリア波形の固定周波数は、対応する高周波数帯域信号の中心周波数にほぼ等しいことを特徴とする請求項58記載の装置。

【請求項60】 対応する高周波数帯域信号を表す再構成波形にノイズ信号を加える加算器をさらに有することを特徴とする請求項51記載の装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、オーディオ符号化方式の分野に関し、特に、人間聴覚系の性質に基づいて高い効率でオーディオ符号化・復号を実行する方法および装置に関する。

【0002】

【従来の技術】例えば音声、音楽などを表すオーディオ信号の蓄積または伝送のいずれかの目的で通常利用される信号を符号化する際に用いられるオーディオ符号化の作業は、長年にわたって注目されている。これは、対応する復号器によって再構成される出力信号の品質を維持しながら、与えられた入力信号を符号化するのに必要なビット数を最小にするためである。例として、音声信号の帯域幅が例えば4kHzである狭帯域音声の場合、最も有効な音声符号器は、当業者に周知の符号励振線形予測(CELP)法に基づいている。このようなCELP法を用いた符号器は通常、4~16kb/sの範囲で動作する。しかし、音源(すなわち、入力音声信号)に関する固有の仮定の結果として、このような符号器は通常、例えば音楽系オーディオ信号に対してはかなり性能が低い。

【0003】他方、知覚オーディオ符号器(PAC)

(例えば、D. Sinha et al., "The Perceptual Audio Coder (PAC)", The Digital Signal Processing Handbook (V. Madisetti and D. Williams, eds.), pp.42-1:42-17, CRC Press, 1998, に記載)のような、知覚基準を用いたオーディオ符号器は、より高いビットレートで、より広帯域のオーディオ信号に対して、かなり良好に動作する。PACのような知覚オーディオ符号器もまた、当業者に周知である(例えば、米国特許第5,040,217号も参照)。具体的には、このような知覚符号器は、音響心理学的モデル(すなわち、人間知覚系の性能に基づくモデル)を用いてスペクトル(すなわち、周波数領域)係数のセットを量子化することにより、聴取者が最終的に区別できないレベルの量子化精度を達成する

のに用いられる符号化ビットの「浪費」を避けることによって、性能の改善を達成する。しかし、残念ながら、約2.4 kb/sより低いビットレートでは、PACのような符号器は通常、音声系オーディオ信号に対してはうまく動作しない。

【0004】最近、マルチモード変換予測符号器(MT-PC: Multimode Transform Predictive Coder) (例えば、S. Ramprasad, "A Multimode Transform Predictive Coder (MTPC) for Speech and Audio", IEEE Speech Coding Workshop, pp.10-12, 1999, に記載)のようなハイブリッド符号器は、上記の両方の符号化パラダイムを組み合わせようと試み、音声およびオーディオの両方の信号に対して、1.6~2.4 kb/sの範囲で相当良好に動作する。しかし、このようなハイブリッド方式の複雑さ(計算量)はしばしば必然的に高くなる。このような方式は、本質的に、2つの異なる技術を単一の符号器に組み合わせなければならないからである。

【0005】符号化オーディオ信号がパケット交換網を通じて伝送されるときにオーディオ符号器に関して必ず生じるもう1つの問題は、パケット損失の問題である。明らかに、このようなアプリケーションの場合、これらの符号器はすべて、このような環境下における有用なツールとなるためには、パケット損失に対してかなりロバストである必要がある。当業者に周知の従来の誤り軽減技術はかなり有効であるが、ほとんどのこのような技術は、約3%より高いパケット損失率では効率が落ちることになる。具体的には、パケット交換網における誤り軽減に対する従来のアプローチは、比較的大きい受信バッファを設けることを含むが、これにより、後に到着するパケットに対する影響が低減される。しかし、双方向通信アプリケーションで使用すると、遅延増大は、双方向通信の有効性に大きい影響を及ぼし、必要以上に高度なエコー制御を必要とすることになる。

【0006】とりわけこのような理由により、当業者に周知のマルチデスクリプティブ(multi-descriptive)符号化技術(例えば、K. Wolf et al., "Source Coding for Multiple Descriptions", Bell System Technical Journal, vol.59, no.9, pp.1417-1426, 1980, に記載)が、この問題点に対する解決法として提案されている。具体的には、この技術によれば、ソース(情報源)符号器は、その情報を複数のほぼ等しい重要性のあるビットストリームに分割する。例えば、2個のストリームの場合、各ストリームは、復号器によって独立に使用されるときには相当の品質レベルを提供するが、両方のストリームが使用されるときには単一ストリームを用いて達成される品質レベルより良好な品質レベルを復号器が提供するように、符号化されることが可能である。パケット損失確率が独立になるように両方のストリームを送信可能であると仮定すると、小さい品質劣化だけで、非常に高い損失率も許容可能となる。

【0007】残念ながら、実際には一般に、符号化効率を低下させずに、マルチデスクリプティブ性を有するオーディオ符号器を設計することはかなり困難であることがわかっている。1つの有効なアプローチは、2チャンネルステレオ音源(例えば、ステレオ音楽)の符号化に関するものである。この場合、音源の固有のマルチデスクリプティブ性(すなわち、2つの独立のチャンネル)を利用して、独立のビットストリームを得ることができる。

【0008】

【発明が解決しようとする課題】上記の理由から、これらの両方の問題点を適切に解決するオーディオ符号化方式を提供することが必要とされている。具体的には、低ビットレートの音声信号と、高ビットレートの音楽信号との両方に対して、ハイブリッド法の複雑さを被らずに、良好に動作する単一のソース符号器/復号器を提供することができれば有利である。さらに、このようなソース符号化方式が、パケット交換環境で用いられるときに誤り軽減の目的で信号の自然なマルチデスクリプティブ分解を提供することができれば有効である。

【0009】

【課題を解決するための手段】本発明の原理によれば、人間聴覚系のいくつかの性質を利用して、音声および音楽の両方の信号に対して良好に動作する単一の非ハイブリッド型オーディオ符号化方式が実現され、さらに、自然なマルチデスクリプティブ分解が得られることが認識された。具体的には、高い周波数では、人間聴覚系は、与えられた基礎となる音響波形刺激に対する同期を失い、この周波数についての時間情報は、聴神経の神経発火の瞬時平均レートによってのみ伝達されることが認識された。

【0010】この認識に基づいて、本発明によれば、符号化されるべきソース信号は複数の周波数帯域に分割され、与えられたしきい値周波数より低い周波数については、波形は、その位相情報を保存するように符号化されるが、そのしきい値より高い周波数については、これらの(高周波側の)それぞれの周波数帯域の波形は、キャリア信号とエンベロープ信号に分解され、(キャリア信号の位相情報ではなく)エンベロープ信号のみが符号化される。その後、復号器では、これらの符号化されたエンベロープ信号が復号され、対応するキャリア波形を変調するために用いられる。キャリア波形は、例えば、周波数帯域の中心周波数に対応する周波数を有する一定の余弦キャリアである。本発明の一実施例によれば、エンベロープ信号は、符号化の前に平滑化(すなわち、ローパスフィルタリング)されることも可能である。

【0011】さらに具体的には、本発明は、オーディオソース信号を符号化する方法および装置であって、

(a) オーディオ信号を、与えられたしきい値周波数より低い周波数成分を含む低周波数帯域信号と、該与えられたしきい値周波数より高い周波数成分を含む高周波数

帯域信号とを含む複数の周波数帯域信号とに分割するステップまたは手段と、(b)前記低周波数帯域信号を表す波形に含まれる位相情報の少なくとも一部を保存するように前記低周波数帯域信号の少なくとも1つを符号化するステップまたは手段と、(c)前記高周波数帯域信号の少なくとも1つに対して、該高周波数帯域信号を表す波形のエンベロープの少なくとも一部を表すが該高周波数帯域信号を表す波形に関連する位相情報をほとんど含まない、対応する臨界帯域エンベロープ信号を生成するステップまたは手段と、(d)対応する臨界帯域エンベロープ信号を符号化することによって前記高周波数帯域信号の少なくとも一部を符号化するステップまたは手段とを有する。

【0012】また、これに対応して、本発明は、このように符号化されたオーディオソース信号を復号する方法および装置であって、(i)それぞれの符号化された低周波数帯域信号を復号して、対応する低周波数帯域信号を表す再構成波形を生成するステップまたは手段と、

(ii)それぞれの符号化された高周波数帯域信号を復号して、該高周波数帯域信号に対応する再構成臨界帯域エンベロープ信号を生成するステップまたは手段と、

(iii)前記再構成臨界帯域エンベロープ信号のそれぞれを、対応するキャリア波形と結合して、対応する高周波数帯域信号を表す再構成波形を生成するステップまたは手段と、(iv)前記対応する低周波数帯域信号を表す再構成波形と前記対応する高周波数帯域信号を表す再構成波形とをそれぞれ結合して、符号化されたオーディオ信号を表す再構成波形を生成するステップまたは手段とを有する。

【0013】さらに、本発明の一実施例によれば、2つの独立のビットストリームが、本発明の符号器によって生成される。一方は、しきい値より高い周波数帯域の適当なサブセットであり、他方は、その適当なサブセットから除かれた他の周波数帯域を表す。(しきい値より低い符号化周波数帯域は、両方のビットストリームに含めることが可能である。)例えば、周波数帯域のこれらの2つのサブセットは、しきい値より高い周波数帯域の系列から、1つおきの帯域をインタリーブする(例えば、1つおきの「偶」と「奇」のサブバンドをインタリーブして「偶」チャンネルおよび「奇」チャンネルを生成するというように)ことによって定義される。このように生成された2つの独立のビットストリームは、それ以外の点*

$$s_i(t) = s(t) * h_i(t) = a_i(t) \cos \phi_i(t) \quad (1)$$

ただし、 $s(t)$ は入力信号であり、 $h_i(t)$ は周波数 $T_i > 1200\text{ Hz}$ を中心とする蝸牛フィルタ i のインパルス応答であり、演算子 $*$ は畳込みを表し、 $a_i(t)$ および $\cos \phi_i(t)$ はそれぞれ、蝸牛信号 $s_i(t)$ のエンベロープおよびキャリア情報である。なお、上記のIHC制限のため、 T_i を起点とするAN

*では従来通りのマルチデスクリプティブ方式で符号化され、最終的に、対応する実施例の復号器によって生成される2チャンネルの再構成信号が得られ、これがそれぞれ左チャンネルおよび右チャンネルとして「再生」されることにより、人間聴覚系のバイノーラル(双聴覚)性を利用することが可能となる。

【0014】

【発明の実施の形態】蝸牛フィルタの出力から時間エンベロープと「キャリア」への分解が、音声近くにおける聴覚メカニズムの役割を定量化するために用いられることは、広く受け入れられている(例えば、J. L. Flanagan, "Parametric Coding of Speech Spectra", Journal of the Acoustical Society of America, vol.68, no. 2, pp.414-430, 1980, を参照)。これは、聴覚系(特に、末梢)がどのように作用するかについての現在のわれわれの理解によって支持される。

【0015】測定されたネコの聴神経(AN: auditory nerve) 応答(例えば、B. Delgutte et al., "Speech Coding in the Auditory Nerve: I. Vowel-like Sounds", Journal of the Acoustical Society of America, vol.75, no.3, pp.866-878, 1984, を参照)と、人間の可能なAN応答との間のアナロジーにより、低CFと高CFの神経繊維の発火パターンの性質間の有意な差が予想される。(当業者に周知のように、CF(特性周波数: Characteristic Frequency)は、基底膜に沿った神経繊維の起点の位置を周波数単位で示す。)低いCFでは、AN神経繊維の神経放電は、基礎となる駆動蝸牛信号に位相ロックされる(すなわち、同期は維持される)。高いCFでは、神経放電の同期は大幅に低下する。実際、このようなCFでは、時間情報は、神経発火の瞬時平均レートによって保持され、これは、基礎となる駆動蝸牛信号の時間エンベロープに関連している。明らかに、これらのAN領域の間に明確な境界はない。むしろ、性質の変化は緩やかである。しかし、本発明の一実施例によれば、この遷移領域は約1200 Hzにあるという作業仮説を採用している。

【0016】このため、本発明の実施例によれば、1200 Hzより高い周波数帯域については、(AN神経発火における同期の喪失によって反映される)キャリア情報を追跡するために、内毛細胞(IHC: Inner Hair Cell)の生理学的限界を利用する。数学的には、次のようにおく。

神経繊維の神経発火は、エンベロープ情報 $a_i(t)$ のみを示し、キャリア情報は失われる。ここで、次の合成信号を考える。

$$\hat{s}_i(t) = a_i(t) \cos \omega_i t \quad (2)$$

【0017】すなわち、もとのキャリア $\cos \phi_i(t)$ が固定余弦キャリア $\cos \omega_i t$ で置き換えら

れている。したがって、帯域制限エンベロープ $a_i(t)$ に対して、 $\hat{s}_i(t)$ は、周波数 ω_i を中心とする帯域制限信号となる。 $\hat{s}_i(t)$ が聴取者の耳に提示された場合、その結果として、蝸牛区画に沿った*

$$\hat{s}(t) = \sum_{i=1}^N \hat{s}_i(t) = \sum_{i=1}^N a_i(t) \cos \omega_i t \quad (3)$$

ただし、 $a_i(t)$ ($i=1, \dots, N$) は、1200 Hz より高い臨界帯域スケールの部分に沿って等間隔 (1個の臨界帯域を1つの間隔とする) に位置するN個の蝸牛フィルタのエンベロープ信号である。(当業者には理解されるように、4 kHz の帯域幅の入力信号に対して、1200 Hz より高い臨界帯域の数は $N=10$ であり、8 kHz の帯域幅の入力信号に対して、対応する値は $N=17$ である。) 多数の有限個の高度に重畳した蝸牛フィルタ (連続する蝸牛区画に沿ったIHCの離散分布によって決定される) によって情報はANに伝達されることを想起すると、もとの信号 $s(t)$ についてのANレベルでの全エンベロープ情報は、IHCの密度によって決定される周波数分解能で表現される。しかし、式(3)によれば、再構成信号 $\hat{s}(t)$ を構成するエンベロープ信号 $a_i(t)$ ($i=1, \dots, N$) は、ANレベルでの全エンベロープ情報の疎なサンプルのみを表現する。

※ 【数2】

$$\hat{s}_{odd}(t) = \sum_{i=1}^{N-1} a_i(t) \cos \omega_i t \quad (4)$$

$$\hat{s}_{even}(t) = \sum_{i=even}^N a_i(t) \cos \omega_i t \quad (5)$$

である(ここで、Nは偶数であると仮定する)。これらのそれぞれの信号における2つの連続する余弦キャリアの間の距離が大きくなることにより、キャリアうなりによる歪みが低減される。 $\hat{s}_{s.a.a.}(t)$ および $\hat{s}_{s.v.e.n.}(t)$ がそれぞれ左耳および右耳に提示されると、聴覚系は、単一の融合した像を生成することになる。

【0020】注目すべき点が2つある。第1に、式(1)の $h_i(t)$ は蝸牛フィルタであり、これは、例えば、当業者に周知のガンマトーンフィルタとして実現される(例えば、B. Gold et al., "Gamma-Tone Filters, Roex Filters, and Auditory Models", Speech and Audio Signal Processing, section 19.4, pp.264-266, John Wiley and Sons, 2000, を参照)。これは、 $h(t) = \sum_{i=1}^N h_i(t)$ がオールパスフィルタでないことを意味する。すなわち、信号 $\sum_{i=1}^N s_i(t)$ (ここで、 $s_i(t)$ は式(1)の変更されていない蝸牛信号である) は、実際には、式(1)のもとの信号 $s(t)$ とは異なる。(しかし、他方では、当業

* 適当な位置(周波数 ω_i に対応する)におけるエンベロープ信号は $a_i(t)$ となる。ここで、次式のようにおく。

【数1】

※【0018】次に、式(3)の $\hat{s}(t)$ を、聴取者の耳に提示する。周波数 ω_i に位置する聴取者の蝸牛フィルタの出力におけるエンベロープは(理想的には)、それぞれの i ($i=1, \dots, N$) に対して、 $a_i(t)$ である。しかし、2つの連続する余弦キャリア周波数 ω_i と ω_{i+1} の間に位置する蝸牛フィルタの出力は、そのフィルタを通過する2つの変調された余弦キャリア信号の「うなり」を反映することになる。これは、好ましくない歪みを生じることがある。

【0019】したがって、本発明の一実施例によれば、うなりによる歪みの可能性を低減するため、臨界帯域をインターリーブしたダイコティック合成(すなわち、左右の耳に提示される信号が異なる場合)を利用する。具体的には、 $\hat{s}_{s.a.a.}(t)$ および $\hat{s}_{s.v.e.n.}(t)$ をそれぞれ、 $\hat{s}(t)$ の奇成分および偶成分の和とする。すなわち、

※ 【数2】

者に周知のように、従来のサブバンド符号化方式におけるフィルタバンクは通常、「完全再構成」要件により設計されることに注意すべきである。)しかし、このような差は不利ではない。その理由は、本発明の原理によれば、われわれはもとの信号を再現しようとしているのではなく、われわれの目的は、もとの信号によって実際に生成された蝸牛エンベロープ情報に対応する、聴取者のANにおける神経活動を刺激する信号を合成することだからである。

【0021】第2に注意すべき点であるが、ここで説明した信号処理技術(すなわち、純粋な余弦キャリアを用いて、基底膜に沿った適当な位置に、サンプリングされたエンベロープ信号を配置すること)は、固有の好ましくない歪みを生じることがある。これは、もとの信号 $s(t)$ が高度に重畳した十分な分解能の蝸牛フィルタバンクを通過すると、その結果得られるエンベロープ情報は、フィルタバンクアレイを通るにつれて緩やかに変化することから理解される。これに対して、式(4)および(5)の $\hat{s}_{s.a.a.}(t)$ および $\hat{s}_{s.v.e.n.}(t)$

を同じフィルタバンクに通すと、エンベロープ情報の変化はずっと粗くなる。式(1)のフィルタ $h_i(t)$ によるエンベロープ情報のサンプリングが疎であるからである。

【0022】上記のように、本発明の原理により得られる情報低減の源の1つは、キャリア情報を追跡する際のIHCの生理学的制限に基づいている。これにより、例えば式(4)および(5)に示したような純粋な余弦キャリアの利用が可能となる。もちろん、符号器により使用される余弦キャリアの周波数は、受信符号化信号を復号する際に用いられる受信機にとって既知である。しかし、上記の解析における蝸牛エンベロープ $a_i(t)$ を、平滑化されたエンベロープ $\tilde{a}_i(t)$ で置き換えることによって、さらに情報を低減することが可能である。この平滑化は、例えばローパスフィルタを用いた従来の方法で実現される。その場合、式(4)および(5)はそれぞれ、次のようになる。

【数3】

$$\tilde{s}_{odd}(t) = \sum_{i \in odd}^{N-1} \tilde{a}_i(t) \cos \omega_i t \quad (6)$$

$$\tilde{s}_{even}(t) = \sum_{i \in even}^N \tilde{a}_i(t) \cos \omega_i t \quad (7)$$

【0023】精神物理学の実験によれば、ローパスフィルタのカットオフ周波数が約250Hzである場合、式(6)および(7)の平滑化された蝸牛エンベロープで合成された音声信号は、式(4)および(5)のものとエンベロープで合成された音声信号と知覚的に区別することができないことが示されている。そこで、本発明の実施例によれば、カットオフ周波数が約250Hzのローパスフィルタを用いて、1200Hzより高い周波数における符号化のための、平滑化されたインタリーブ臨界帯域エンベロープを生成する。その後、このように符号化されたエンベロープは、上記のように受信機で復号され、対応する信号の対が、式(6)および(7)に従って合成される。

【0024】なお、チャンネルをインタリーブしたダイコティック合成を用いることによって、式(6)および(7)に従って合成される $\sim s_{aa}(t)$ および $\sim s_{vva}(t)$ は、1200Hzより上では無関係になる。したがって、本発明のさまざまな実施例によれば、いくつかのマルチデスクリプティブ合成法のうちの任意のものを用いることが可能となる。例えば、受信機において(例えば、測定されるチャンネル損失に依存して)、左耳(L)および右耳(R)には、次のいずれかのように供給される。

(1) $\sim s_{aa}(t)$ をLに、 $\sim s_{vva}(t)$ をRに。

(2) $\sim s_{aa}(t)$ をLおよびRの両方に。

(3) $\sim s_{vva}(t)$ をLおよびRの両方に。

【0025】なお、本発明のアプローチは、場合によっては、次の2種類の人為的産物を生じることがある。第1に、純粋余弦キャリアの利用により、式(4)および(5)の $\hat{s}_{aa}(t)$ および $\hat{s}_{vva}(t)$ に知覚可能な歪みを生じることがあり、その量は、スペクトル内容とキャリア周波数の間の相互作用と、聴取者の経験とに依存する。第2に、ダイコティック提示は、ダイオティック(すなわち、同じ信号が両耳に提示される場合)提示によって生成されるものとは異なる空間像を生成することがある。本発明の実施例の方法をマルチデスクリプティブ方式のアプローチに従って用いると、ダイコティックモードからダイオティックモードへの切換により、像の空間位置の切換が起こる。この問題点は、ステレオヘッドフォンを用いるのではなく、2個のスピーカを通じて2チャンネル出力を行うような実世界のアプリケーション(例えば、デスクトップアプリケーションの場合)では軽減される可能性がある。

【0026】[実施例]図1に、本発明の技術に従って、高周波数において平滑化されたインタリーブ臨界帯域エンベロープを用いてマルチデスクリプティブステレオオーディオ符号化および復号を実行するシステムで用いられる、オーディオ符号器の実施例を示す。なお、具体的には、符号器から復号器へ送信される情報は、符号化されたベースバンド信号(すなわち、1200Hzより低い周波数を含む信号の部分)と、符号化された平滑化(すなわち、ローパスフィルタリングされた)臨界帯域エンベロープとからなる。さらに、注意すべき点であるが、図示した本発明の実施例によれば、2個のビットストリームが生成される。一方は、(ベースバンド信号の符号化表現とともに)「偶」チャンネルの符号化表現を含み、他方は、(同じくベースバンド信号の符号化表現とともに)「奇」チャンネルの符号化表現を含む。

【0027】具体的には、図1の実施例の符号器は、1対の符号化ビットストリーム、すなわち、符号化された偶チャンネルビットストリームおよび符号化された奇チャンネルビットストリームをそれぞれ生成するための、奇チャンネル符号器12-1および偶チャンネル符号器12-2を有する。符号化されるべきオーディオ信号を含む入力信号 $s(t)$ は、奇チャンネル符号器12-1および偶チャンネル符号器12-2の両方に入力され、それぞれの対応するフィルタバンクを通じて送られる。例えば、奇チャンネル符号器12-1のフィルタバンクは、ベースバンド信号(すなわち、1200Hz以下の信号の周波数成分)を生成する1200Hzローパスフィルタ14-1と、与えられた臨界帯域に対応するサブバンド信号を生成する、1200Hzより高いすべての「奇数番号の」臨界帯域のバンドパスフィルタ(すなわち、図示したようなバンドパスフィルタ17-1ないしバンドパスフィ

ルタ17-(N-1))とを有する。同様に、偶チャネル符号器12-2のフィルタバンクは、ベースバンド信号を生成する1200Hzローパスフィルタ14-2と、同じく与えられた臨界帯域に対応するサブバンド信号を生成する、1200Hzより高いすべての「偶数番号の」臨界帯域のバンドパスフィルタ(すなわち、図示したようなバンドパスフィルタ17-2ないしバンドパスフィルタ17-N)とを有する。これらのフィルタバンクの各フィルタは全く従来のものであり、当業者に周知である。バンドパスフィルタのセットは、バークスケールに沿って一様に分布し、各フィルタは、例えば当業者に周知のガンマトーンフィルタのような「蝸牛」周波数応答を有する。(同じく当業者に周知のように、バークスケールは、蝸牛の実際の帯域を表す近似的に對数的な周波数スケールを有する。)なお、ここでは、説明を簡単にするため、N(1200Hzより高い臨界帯域の数)は偶数であると仮定する。

【0028】本発明の原理によれば、(奇チャネル符号器12-1および偶チャネル符号器12-2の両方における)それぞれの高周波数サブバンド信号に対して、エンベロープ生成器18-i(バンドパスフィルタ17-iによって生成されるサブバンド信号に対応する)は、信号のキャリア部分(すなわち、上記の数学的解析による $\cos \phi_i(t)$)を除去することにより、符号化のために、信号の「エンベロープ」部分 $a_i(t)$ のみを残す。次に、これらの生成されたそれぞれの「エンベロープ」信号に対して、250Hzローパスフィルタ19-i(エンベロープ生成器18-iによって生成されたエンベロープ信号に対応する)は、エンベロープ信号を(同じく上記の数学的解析に従って)「平滑化」する。なお、エンベロープ生成器および250Hzローパスフィルタはいずれも従来のものであり、当業者に周知である。最後に、(奇チャネル符号器12-1内の)量子化器16-1および(偶チャネル符号器12-2内の)量子化器16-2は、同じく当業者に周知の従来の量子化技術(例えば、後述の例示的な実装についての説明を参照)を用いて、ベースバンド信号と、個々の臨界帯域信号の平滑化されたエンベロープとを符号化して、対応する符号化ビットストリームを生成する。

【0029】図2に、本発明の技術に従って、高周波数において平滑化されたインターリーブ臨界帯域エンベロープを用いてマルチデスクリプティブステレオオーディオ符号化および復号を実行するシステムで用いられる、図1のオーディオ符号器に対応する復号器の実施例を示す。再び注意すべき点であるが、符号器から復号器へ送信された情報は、符号化されたベースバンド信号(すなわち、1200Hzより低い周波数を含む信号の部分)と、符号化された平滑化(すなわち、ローパスフィルタリングされた)臨界帯域エンベロープとからなる。さらに、再び注意すべき点であるが、図示した本発明の実施

例によれば、2個のビットストリームが復号器によって受信される。一方は、(ベースバンド信号の符号化表現とともに)「偶数番号の」サブバンドの符号化表現を含み、他方は、(同じくベースバンド信号の符号化表現とともに)「奇数番号の」サブバンドの符号化表現を含む。

【0030】具体的には、図2の実施例の復号器は、受信される符号化ビットストリームの対、すなわち、符号化された偶チャネルビットストリームおよび符号化された奇チャネルビットストリームをそれぞれ復号するための、奇チャネル復号器22-1および偶チャネル復号器22-2を有する。まず、(奇チャネル復号器22-1内の)逆量子化器24-1および(偶チャネル復号器22-2内の)逆量子化器24-2は、ビットストリームを復号して、個々の信号、すなわち、(奇チャネル復号器および偶チャネル復号器のそれぞれにおける)ベースバンド信号と、含まれている平滑化された臨界帯域エンベロープ信号のセットとを生成する。具体的には、奇チャネル復号器22-1の逆量子化器24-1は、ベースバンド信号と、奇数番号のそれぞれのサブバンドに対する平滑化された臨界帯域エンベロープ信号とを生成し、一方、偶チャネル復号器22-2の逆量子化器24-2は、ベースバンド信号も生成するが、それとともに、偶数番号のそれぞれのサブバンドに対する平滑化された臨界帯域エンベロープ信号を生成する。逆量子化器24-1および24-2は、それぞれ図1の奇チャネル符号器12-1および偶チャネル符号器12-2の量子化器16-1および16-2の機能に対応する(すなわち、それらの逆の機能を実行する)が、同じく全く従来のものであって、当業者に周知である。

【0031】次に、それぞれの平滑化された臨界帯域エンベロープ信号に対して、奇チャネル復号器22-1および偶チャネル復号器22-2はそれぞれ、乗算器回路25-i(サブバンドiに対する再生された平滑化臨界帯域エンベロープ信号に対応し、これを処理する)を有する。乗算器回路25-iは、固定余弦キャリア、すなわち、 $\cos \omega_i t$ (ただし、 ω_i は、臨界帯域iの中心周波数となるようにに選ぶ)を対応する平滑化臨界帯域エンベロープ信号に乘じる。このようにして、結果として得られる合成信号は、与えられた臨界帯域を中心とする固定余弦キャリア信号を、与えられた臨界帯域に対応する符号化され送信された平滑化臨界帯域エンベロープ信号を復号したもので変調した信号を含む(上記の数学的解析と、特に式(2)およびその説明とを参照)。

【0032】次に、対応する乗算器回路25-iによって生成された、それぞれの再構成された臨界帯域信号に対して、対応する加算器回路26-iは、対応する低強度帯域制限ノイズ信号 $n_i(t)$ を加算する。これは、例えば当業者に周知の白色ガウシアンノイズのような低レベルのノイズを、隣接する未占有の周波数帯域に「充

填」することによって、再構成された信号における人工的ノイズを低減する。具体的には、占有帯域（これにノイズ信号が加算される）内の実際の信号の強度に対する、ノイズ信号の信号対ノイズ比は、約-10dBとなり、ノイズ信号の帯域幅は、3個の連続する周波数帯域（ノイズが加算される周波数帯域を中心とする）をカバーすることが可能である。このような低強度ノイズ信号は従来のものであり、当業者が容易に定義することが可能である。なお、このノイズ信号もまた、蝸牛型にすることも可能である。

【0033】最後に、奇チャネル復号器22-1の加算器回路27-1は、復号されたベースバンド信号と、奇数番号のサブバンドに対するそれぞれの合成されたサブバンド信号（加算器回路26-1, 26-3, ..., 26-(N-1)によって生成されたもの）との和をとり、上記のような、結果として得られる合成信号 $s_{aa}(t)$ を生成する（例えば、式(6)を参照）。同様に、偶チャネル復号器22-2の加算器回路27-2は、復号されたベースバンド信号と、偶数番号のサブバンドに対するそれぞれの合成されたサブバンド信号（加算器回路26-2, 26-4, ..., 26-Nによって生成されたもの）との和をとり、上記のような、結果として得られる合成信号 $s_{vva}(t)$ を生成する（例えば、式(7)を参照）。こうして、本発明の原理によれば、図示されているように、 $s_{aa}(t)$ はスピーカ28-1に送られ、 $s_{vva}(t)$ はスピーカ28-2に送られる（あるいは、別の実施例では、この逆に送られる）ことにより、同じく前述のように、聴取者29の左耳には信号 $s_{aa}(t)$ の音響的再現が供給され、右耳には信号 $s_{vva}(t)$ の音響的再現が供給される。

【0034】[例示的な実現に関する追加的考察]本発明の実施例によれば、本発明で用いるような技術に基づく符号化システムの全体の複雑さ（計算量）および遅延は、ここで説明するように制約される。なお、当業者に周知の追加技術を用いれば、場合により遅延の増大の犠牲を払って、さらに良好な符号化効率を得ることも可能である。

【0035】例えば、図1のバンドパスフィルタバンクは、当業者に周知の128タップFIR（有限インパルス応答）フィルタで実現することにより、8msの遅延が導入される（すなわち、8000サンプル/秒）。さまざまな信号間の直接的な時間関係を維持するために、ベースバンド信号（帯域幅1200Hz）は1/3の比で（すなわち、2666サンプル/秒に）ダウンサンプリングされ、ローパスフィルタリングされたエンベロープ信号（帯域幅250Hz）は1/15の比で（すなわち、533サンプル/秒に）ダウンサンプリングされる（このようなダウンサンプリングは実施例では図示していないが、当業者に周知である）。ダウンサンプリング

された周波数での符号化遅延は、それぞれのダウンサンプリング比だけ遅延を増大させることになる。したがって、この実施例によれば、サンプルごとに作用する符号化方式を用いることが可能である。例えば、デルタ変調やADPCM（いずれも当業者に周知である）のような符号化技術を用いることが可能である。

【0036】上記のダウンサンプリングされたエンベロープ信号は量子化ノイズに対して非常にロバストであることがわかっているため、2ビット/サンプル量子化器による簡単なADPCM法を用いることが可能である。これに対して、ベースバンド信号は、量子化誤りに対して、より敏感であることがわかっている。具体的には、ADPCMを用いると、3ビット/サンプルで、受容可能な品質レベルが得られることがわかっている。本発明の他の実施例によれば、符号器のベースバンド部分を用いてマルチデスクリプティブ符号化方式を実装することが可能であるが、単純化および効率化のために、ここで具体的に説明している本発明の実施例によれば、ベースバンド情報は全部、両方のストリームで用いられる。こうして、例として、この実施例の全ビットレートは次のようになる。

2チャネル×（ベースバンド：8kb/s+エンベロープ：5×1.066kb/s）=26.66kb/s
また、およその全体の符号化の複雑さ（計算量）は例えば次のようになる。

（ベースバンド：1.3MIPS+エンベロープ：10×0.3MIPS）=4.3MIPS

【0037】制限サイズVQ（ベクトル量子化器。当業者に周知である）による後方適応予測（同じく当業者に周知である）は、大きいアルゴリズム遅延を導入せずに、サンプルあたりのビット数をさらに低減することができる。なお、予測器次数は、当業者には容易に判断されるように、スペクトルの正確な記述を生成するために特に高い必要はない。本発明のこの実施例による上記のようなダウンサンプリングレートの特定の選択により、追加の符号化遅延を導入せずに、5次元VQを使用することが可能となる。さらに、本発明の実施例によれば、ダウンサンプリングされた信号に対して16次の予測器を用いて、LD-CELP（低遅延符号励振線形予測。当業者に周知の符号化技術である）の修正版を使用し、顕著な可聴劣化なしで、8MIPSの計算量で、2ビット/サンプル（すなわち、2666サンプル/秒）を得ることが可能である。

【0038】復号器では、信号は、アップサンプリングされ、当業者に周知の32タップFIR（有限インパルス応答）フィルタで補間されて、さらに2msの遅延が生じる。このアプローチの全体の計算量は、フィルタリングおよびアップサンプリングについては約14MIPSであり、量子化については約11MIPSである。全体のエンドツーエンド遅延は、フィルタリングによる1

0msと、符号化による2msである。全ビットレートは例えば次のようになる。

2チャンネル×(ベースバンド: 5.33kb/s+エンベロープ: 5×1.066kb/s)=21.332kb/s

最後に、注意すべき点であるが、広帯域信号(8kHz)の場合、ビットレートの正味の増大は、各ビットストリームごとに3~4kb/sにしかならない。

【0039】[詳細な説明への付記]以上、本発明について詳細に説明したが、当業者であれば、上記の記載に基づいて、さまざまな変形例を考えることが可能である。例えば、当業者には認識されるように、図面のブロック図は、本発明の原理を実現する例示的な回路の概念図を表す。同様に、認識されるように、流れ図、状態遷移図、擬似コードなどは、実質的に、コンピュータ可読媒体内に表現し、コンピュータあるいはプロセッサによって実行されるようにすることが(そのようなコンピュータあるいはプロセッサが明示的に図示されているかどうかにかかわらず)可能である。

【0040】図示されたさまざまな要素の機能は、「プロセッサ」あるいは「モジュール」と名づけられた機能ブロックを含めて、専用ハードウェアや、適当なソフトウェアを実行可能なハードウェアを用いて提供可能である。プロセッサによって提供されるとき、これらの機能は、単一の専用プロセッサによって、単一の共有プロセッサによって、あるいは、複数のプロセッサ(その一部は共用可能)によって提供可能である。さらに、「プロセッサ」あるいは「コントローラ」という用語の明示的な使用は、ソフトウェアを実行可能なハードウェアを限定的に指すものと解釈すべきではなく、暗黙的に、限定的でなく、デジタル信号プロセッサ(DSP)ハードウェア、ソフトウェアを記憶するための読み出し専用メモリ(ROM)、ランダムアクセスメモリ(RAM)、および不揮発性記憶装置を含む。他のハードウェアも、通常のものであるカスタム化したものにかかわらず、含まれる。同様に、図中にスイッチがあれば、それは概念的なものにすぎない。その機能は、プログラムロジックの動作によって、専用ロジックによって、プログラム制御と専用ロジックの相互作用によって、あるいは、手動でも、実行可能であり、具体的には文脈から理

能である。

【0041】特許請求の範囲において、特定の機能を実行する手段として表現される要素は、例えば、(a)その機能を実行する回路要素の組合せ、あるいは、(b)ファームウェア、マイクロコードなどを含む任意の形のソフトウェアを、その機能を実行するようにそのソフトウェアを動作させるための適当な回路と組み合わせたもの、を含めて、その機能を実行する任意の態様を含む。

【0042】

【発明の効果】以上述べたごとく、本発明によれば、人間聴覚系のいくつかの性質を利用して、音声および音楽の両方の信号に対して良好に動作する単一の非ハイブリッド型オーディオ符号化方式が実現され、さらに、自然なマルチデスクリプティブ分解が得られる。

【図面の簡単な説明】

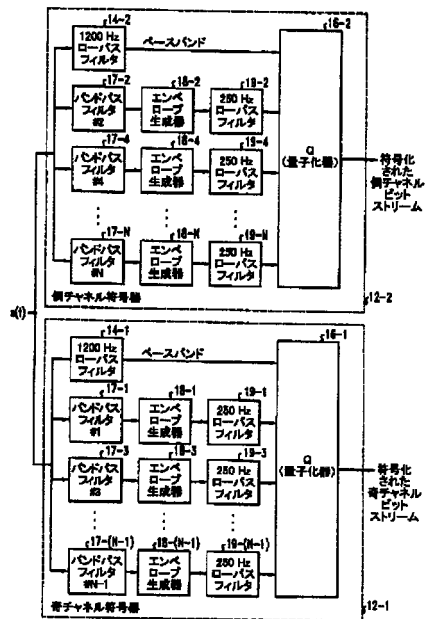
【図1】本発明の技術に従って、高周波数において平滑化されたインタリーブ臨界帯域エンベロープを用いてマルチデスクリプティブステレオオーディオ符号化および復号を実行するシステムで用いられる、オーディオ符号器の実施例の図である。

【図2】本発明の技術に従って、高周波数において平滑化されたインタリーブ臨界帯域エンベロープを用いてマルチデスクリプティブステレオオーディオ符号化および復号を実行するシステムで用いられる、図1のオーディオ符号器に対応する復号器の実施例の図である。

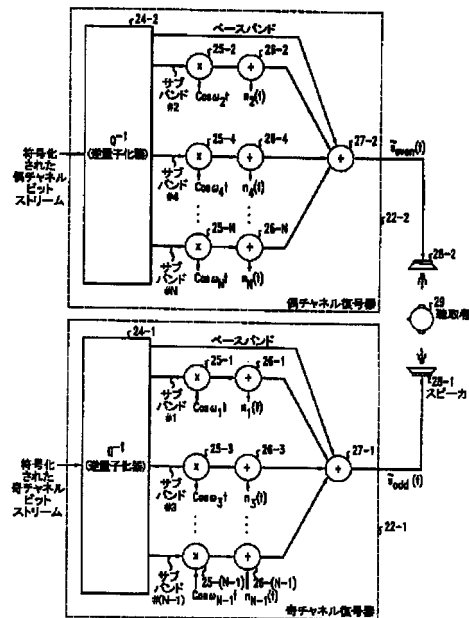
【符号の説明】

- 12-1 奇チャンネル符号器
- 12-2 偶チャンネル符号器
- 14 1200Hzローパスフィルタ
- 16 量子化器
- 17 バンドパスフィルタ
- 18 エンベロープ生成器
- 19 250Hzローパスフィルタ
- 22-1 奇チャンネル復号器
- 22-2 偶チャンネル復号器
- 24 逆量子化器
- 25 乗算器回路
- 26, 27 加算器回路
- 28 スピーカ
- 29 聴取者

【図1】



【図2】



フロントページの続き

(71)出願人 596077259

600 Mountain Avenue,
Murray Hill, New Jersey
07974-0636 U. S. A.

(72)発明者 オデッド ギザ

アメリカ合衆国、07090 ニュージャージー
州、ウェストフィールド、アーチボール
ド パレス 128

Fターム(参考) 5D045 DA11

5J064 AA02 BA05 BA17 BB03 BC02
BC12 BC16 BC18 BC21 BC22
BD02